

Generalized Bell Inequality Experiments and Computation

Matty J. Hoban,^{1,2} Joel J. Wallman,³ and Dan E. Browne¹

¹*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom.*

²*Department of Computer Science, University of Oxford,*

Wolfson Building, Parks Road, Oxford OX1 3QD, United Kingdom.

³*School of Physics, The University of Sydney, Sydney, New South Wales 2006, Australia.*

(Dated: November 24, 2011)

We consider general settings of Bell inequality experiments with many parties, where each party chooses from a finite number of measurement settings each with a finite number of outcomes. We investigate the constraints that Bell inequalities place upon the correlations possible in a local hidden variable theories using a geometrical picture of correlations. We show that local hidden variable theories can be characterized in terms of limited computational expressiveness, which allows us to characterize families of Bell inequalities. The limited computational expressiveness for many settings (each with many outcomes) generalizes previous results about the many-party situation each with a choice of two possible measurements (each with two outcomes). Using this computational picture we present generalizations of the Popescu-Rohrlich non-local box for many parties and non-binary inputs and outputs at each site. Finally, we comment on the effect of pre-processing on measurement data in our generalized setting and show that it becomes problematic outside of the binary setting, in that it allows local hidden variable theories to simulate maximally non-local correlations such as those of these generalised Popescu-Rohrlich non-local boxes.

I. INTRODUCTION

Quantum mechanics is incompatible with a classical view of the world in many ways. In particular, quantum theory is incompatible with the assumption of local realism as there are entangled quantum states and measurements that lead to violations of a Bell inequality [1]. Consequently, Bell inequalities provide an extremely clear distinction between classical, local hidden variable (LHV), and non-classical (e.g., quantum) theories and so have been studied with great fervor since their discovery.

Another motivation for the study of Bell inequalities has been from an informational point-of-view. Quantum information has brought new insight into the very nature of quantum mechanics and Bell inequalities have been used to gain insight into the information processing power of quantum mechanics. Bell inequality violations have been used to guarantee the security of quantum key distribution [2], generate randomness [3], give an advantage in communication complexity [4] and non-local games [5]. Recently, connections have been made between quantum computing and a violation of a Bell inequality as giving some computational advantage [6–8].

Despite their utility, there are many open questions about Bell inequalities. Many breakthroughs have been made by constructing Bell inequalities for many parties each with a choice of two measurements and two measurement outcomes [9–11]. However, constructing Bell inequalities (that fully define LHV correlations) in general is NP-hard in the number of measurement choices and measurement outcomes at each site [12].

Generalizations of the Bell inequality experiment away from the two measurement setting, two measurement outcome scenario have been studied and have led to the discovery of interesting phenomena [13–18]. For example, quantum violations can be greater [14] and more robust to experimental imperfections [19] in these general scenarios.

The Clauser-Horne (CH) inequality [20] also showed that

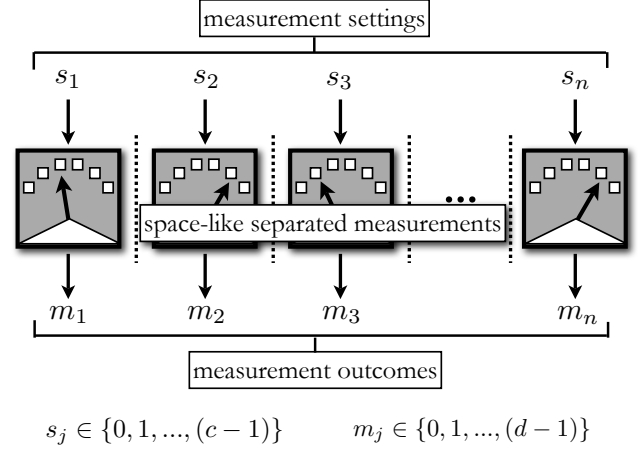


FIG. 1: In a single run of a general Bell experiment, n parties each make a measurement from c possible choices, where each measurement has d possible outcomes. Labelling the j th parties' measurement choice and outcome by s_j and m_j respectively, we can describe each run of the experiment with n -digit strings \mathbf{m} and \mathbf{s} .

the full probability distribution (including marginal probabilities) can be constrained in LHV theories and violated by quantum mechanics. However, quantum theory does not obtain the maximal violation of the CH inequality possible in theories that do not allow superluminal signaling (non-signaling for short), as demonstrated by Popescu-Rohrlich boxes [21]. As a result, much work has gone into characterizing the space of non-signaling probability distributions [22]. An approach based around convex polytopes has been particularly fruitful but the space of all non-signaling probability distributions remains complicated and often counter-intuitive, as demonstrated by the Guess Your Neighbor's Input non-local game [23].

The convex polytope of non-signaling probability distributions in the CH inequality setting is well-studied and is an excellent platform for describing quantum correlations via an information theoretic, or physical principle [24–27]. For more general settings, it is difficult to describe the space of non-signaling probabilities. To avoid this difficulty yet still obtain some intuition about the space of non-signaling probabilities, we consider the space of correlators, i.e., the space of probability distributions over joint outcomes.

For two-setting, two-outcome Bell inequalities, the correlations studied in Bell experiments can be described in terms of stochastic Boolean maps, e.g. probabilistic mixtures of Boolean functions from the two bits describing the measurement settings to the parity of the measurement outcomes, which is a single bit. This approach has proven well-suited to the study of many-body generalizations of the Clauser-Horne-Shimony-Holt (CHSH) inequality and led to a number of unifications and new insights [6–8].

In this paper, we generalize this approach by replacing stochastic Boolean maps with stochastic maps from the set of measurement settings to a single digit representing the parity of the output digits. Our generalization reveals new phenomena which do not occur in the simpler two-setting, two-outcome case.

Hand-in-hand with this “computational” description of correlations we shall use a geometric approach. The list of conditional probabilities describing the stochastic map can also be treated as a vector in a real vector space [12, 28, 29]. In this picture, the valid sets of correlations which occur in a theory form a convex region in this space, which, in the case of LHV theories, is a polytope. A polytope can be defined as the intersection of the half-spaces that satisfy a set of linear inequalities, which, for the LHV polytope, are simply the facet Bell inequalities.

Very recently, two of us showed that for 2-setting 2-outcome Bell inequalities, computational expressiveness provides an elegant method for studying Measurement-based Quantum Computing [7, 8]. We shall show how such approaches can be generalized, and that the switch from Boolean to more general maps allows more complex behavior than in the binary case.

In Section II, we outline the general framework, fix notation and introduce the modular arithmetic which will be employed throughout the paper. In Section III we construct Bell inequalities in this framework. In Section IV we study how methods from [7] and [8] can be generalized and then summarize in Section V.

II. FRAMEWORK FOR GENERAL CHSH TESTS

The general scenario we consider is the n -partite Bell experiment shown in Fig. 1, where $n > 1$. For simplicity, we assume all space-like separated parties choose from c possible measurement settings, each with d possible outcomes, so we label an experiment by the variables (n, c, d) .

As in the standard CHSH experiment, we assume that each party’s measurement setting is chosen randomly from a uni-

form distribution and is uncorrelated with the state of the system (see [30–33] for some consequences of dropping this assumption).

We label the c possible measurements and d possible measurement outcomes by digits $s_j \in \{0, \dots, c-1\} \in \mathbb{Z}_c$ and $m_j \in \mathbb{Z}_d$ respectively. Therefore the n settings and measurements for each repetition of the experiment are labeled by n digit-strings $\mathbf{m} \in \mathbb{Z}_d^n$ and $\mathbf{s} \in \mathbb{Z}_c^n$ corresponding to all n measurement outcomes and all n settings respectively, where \mathbf{x} denotes the n -digit string such that the j -th digit is x_j .

To simplify the analysis we shall assume initially that both c and d are prime. We discuss features of the analysis in the non-prime case in appendix B. We will employ arithmetic in both modulo c and modulo d . As we are considering correlators obtained by adding measurement outcomes, the majority of the arithmetic will be modulo d . Therefore we use the following convention. In all expressions, arithmetic will be modulo d , except in exceptional cases where the arithmetic is modulo c , which will be denoted with large square brackets and subscript c , i.e. $[\dots]_c$.

After the data \mathbf{m} and \mathbf{s} has been collected from many repetitions of the experiments, conditional probabilities $p(\mathbf{m}|\mathbf{s})$ for each value of \mathbf{m} and \mathbf{s} can be calculated. In the initial CHSH paper, and in many-body generalizations, rather than studying this full probability distribution, one merely considers the statistics of the parity of the output \mathbf{m} . Here, we shall take a similar course and study correlators

$$p(k|\mathbf{s}) = p\left(\sum_{j=1}^n m_j = k|\mathbf{s}\right), \quad (1)$$

where $p(\sum_{j=1}^n m_j = k|\mathbf{s}) = \sum_{\mathbf{m}|\sum_{j=1}^n m_j = k} p(\mathbf{m}|\mathbf{s})$. When taking a sum over m_j we are performing addition modulo d but when taking sums of probabilities or describing anything that is not related to measurement settings or outcomes we are performing standard, natural arithmetic over the reals. This natural generalization of the CHSH-type correlators in (1) also encompasses the well-known Collins-Gisin-Linden-Masser-Popescu (CGLMP) inequality for two parties [14]. These correlators have a well-defined role when considering the expectation values of joint measurements [35]. Moreover, singling out this one parameter simplifies the problem by reducing the dimensionality of the problem, whilst still capturing interesting properties of the full distribution $p(\mathbf{m}|\mathbf{s})$.

In Section II A we introduce the mathematical framework for describing these correlators, which employs a correspondence between conditional probabilities and stochastic maps. In Section II B we apply this framework and derive the full family of correlators that are possible in LHV theories. Then, in Section II C we consider correlators in more general non-signaling theories and derive generalizations of the Popescu-Rohrlich non-local box [21].

A. Correlators as Stochastic Maps

Before we begin our study of correlators, we need to introduce some notation and mathematical construction. The correlators

$p(k|\mathbf{s})$ are maps from n digit-strings $\mathbf{s} \in \mathbb{Z}_c^n$ to probability distributions over a single digit $k \in \mathbb{Z}_d$. Any stochastic map is a convex combination of deterministic processes, which in this case are functions $f : \mathbb{Z}_c^n \rightarrow \mathbb{Z}_d$.

Such functions take an “input” $\mathbf{s} \in \mathbb{Z}_c^n$ to a single “output” $k \in \mathbb{Z}_d$. Since \mathbb{Z}_d and \mathbb{Z}_c^n are fields for c, d prime, these functions can be represented as elements of a vector space. There are many bases one could choose to represent the function. Perhaps the simplest and cleanest basis is the set of Kronecker delta functions $\delta_{\mathbf{s}}^{\mathbf{y}}$, where $\mathbf{y} \in \mathbb{Z}_c^n$, which equals 1 when $\mathbf{y} = \mathbf{s}$ and 0 otherwise. Any function can be represented as

$$f(\mathbf{s}) = \sum_{\mathbf{y} \in \mathbb{Z}_c^n} f(\mathbf{y}) \delta_{\mathbf{s}}^{\mathbf{y}}. \quad (2)$$

The d^{c^n} possible coefficients correspond to d^{c^n} functions. It will sometimes be convenient for the basis set to include the constant function. We shall then replace the delta function $\delta_{\mathbf{s}}^{\mathbf{0}}$ where $\mathbf{0}_c$ is the all-zeros string with the constant function to write

$$f(\mathbf{s}) = \left(\sum_{\mathbf{y} \in (\mathbb{Z}_c^n - \mathbf{0}_c)} \epsilon_{\mathbf{y}} \delta_{\mathbf{s}}^{\mathbf{y}} \right) + \alpha, \quad (3)$$

where $\epsilon_{\mathbf{y}} = (f(\mathbf{y}) - \alpha)$ and $\alpha \in \mathbb{Z}_d$.

The natural expression for the computational power of correlators in a general theory is how close they can approximate an arbitrary function. Therefore we want to be able to rewrite (3) as a polynomial. This can be done as the Kronecker delta $\delta_{\mathbf{s}}^{\mathbf{y}}$ can be written as a polynomial (modulo c),

$$\begin{aligned} \delta_{\mathbf{s}}^{\mathbf{y}} &= \prod_{j=1}^n \delta_{y_j}^{s_j} = \prod_{j=1}^n \left[1 - (s_j - y_j)^{(c-1)} \right]_c, \\ &= \prod_{j=1}^n \left[1 - \sum_{l=0}^{(c-1)} (-1)^l \binom{c-1}{l} (y_j)^l (s_j)^{c-(l+1)} \right]_c. \end{aligned} \quad (4)$$

The first line follows from Fermat’s little theorem, as $a^{b-1} \equiv 1 \pmod{b}$ if a is non-zero and b is prime. To reach the second line we have used the binomial theorem which follows from the distributivity of modular arithmetic. This shows us that any function $f(\mathbf{s})$ can be expressed as a polynomial with mixed modular arithmetic systems. Furthermore, when $d \geq c$ we can write the function solely in terms of modulo d arithmetic since equation (4) holds for any prime $\geq c$.

In this paper, the following class of functions, which we term the “ n -partite linear functions”, will be crucial in the study of LHV theories.

Definition 1. A function $f(\mathbf{s}) : \mathbb{Z}_c^n \rightarrow \mathbb{Z}_d$ is an **n -partite linear function** if it can be expressed as

$$f(\mathbf{s}) = \sum_{j=1}^n g_j(s_j), \quad (5)$$

where addition is modulo d and the $g_j(s_j)$ are functions $\mathbb{Z}_c \rightarrow \mathbb{Z}_d$ of a single variable. If $f(\mathbf{s})$ is not an n -partite linear function, we refer to it as a **non- n -partite linear function**.

These functions will play a role analogous to that of the linear Boolean functions in [7, 8]. The n -partite linear functions can also be written as

$$f(\mathbf{s}) = \alpha + \sum_{j=1}^n \sum_{a=1}^{(c-1)} \beta_{j,a} \delta_a^{s_j}, \quad (6)$$

with $\alpha, \beta_{j,a} \in \mathbb{Z}_d$. There are $d^{n(c-1)+1}$ n -partite linear functions. For c being prime, the key feature of a non- n -partite linear function is the presence of cross-multiplicative terms in (4) between different digits s_j of \mathbf{s} .

For non-prime c , we cannot write delta functions in the neat polynomial of (4) but in a more elaborate fashion (see appendix B) but we can still work in terms of the delta functions for non-prime c . Therefore, the definition of an n -partite linear function in (6) applies for non-prime c . Thus for prime and non-prime c , the delta functions $\delta_a^{s_j}$ only singularly define a value of $f(\mathbf{s})$ when there is only one non-zero element in \mathbf{s} . For other values of \mathbf{s} there is addition modulo d between the terms $\delta_a^{s_j}$. Equivalently then, for $f(\mathbf{s})$ to be a non- n -partite linear function, there must be at least one delta function $\delta_{\mathbf{y}}^{\mathbf{s}}$ in 3 with \mathbf{y} being a digit-string with more than one non-zero element. A function $f(\mathbf{s})$ can be written in terms of an n -partite linear part plus a non- n -partite linear part (with addition modulo d), and if the n -partite linear part is zero, then $f(\mathbf{s}) = 0$ for all \mathbf{s} with only one non-zero element.

Finally, we note that the above treatment also applies in a more general setting where instead of setting digit-strings $\mathbf{s} \in \mathbb{Z}_c^n$ we now have digit-strings $\mathbf{s} \in \mathbb{Z}_{c_1} \times \mathbb{Z}_{c_2} \times \dots \times \mathbb{Z}_{c_n} = \bigoplus_{j=1}^n \mathbb{Z}_{c_j}$. That is, the number of possible settings, now labeled c_j at each j th site could now be different from other sites and also non-prime. The string \mathbf{s} is now a direct sum (or Cartesian product) of the registers \mathbb{Z}_{c_j} for each site. All delta functions as defined can be carried over to this general setting. Thus single site maps and n -partite linear functions are exactly the same as (6) but for the j th site, c is replaced with c_j . We now discuss a geometrical picture of the correlators $p(k|\mathbf{s})$.

The correlators $p(k|\mathbf{s})$ are stochastic maps and so must satisfy the positivity inequalities, i.e., the set of $c^n(d-1)$ linear inequalities $p(k|\mathbf{s}) \geq 0$, for all k, \mathbf{s} . Furthermore, as some outcome always occurs for any choice of measurement settings \mathbf{s} , the correlators must also satisfy the normalization equations, $\sum_{k \in \mathbb{Z}_d} p(k|\mathbf{s}) = 1$ for all \mathbf{s} . We then only have $(d-1)$ independent correlators for each value of \mathbf{s} , so we omit the correlator $p(0|\mathbf{s})$ and treat the remaining correlators as elements of a vector in a $c^n(d-1)$ -dimensional real space with elements being $p(k|\mathbf{s})$ for $k \neq 0$.

The space of correlators satisfying the linear inequalities of the positivity conditions $p(k|\mathbf{s}) \geq 0$ and normalisation conditions in this reduced space, $\sum_{k \neq 0} p(k|\mathbf{s}) \leq 1$, is then a convex polytope labeled \mathcal{P} . It is a convex polytope as it is the intersection of the half-spaces defined by these linear inequalities. In any specific theory, the region of allowed correlators, \mathcal{T} , will be a convex subregion of \mathcal{P} . Equivalently, \mathcal{P} can be defined as the convex hull of all deterministic correlators $p(k|\mathbf{s}) \in \{0, 1\}$, for all k and \mathbf{s} ; these deterministic correlators are the extreme points of \mathcal{P} .

If \mathcal{T} has a finite number of extreme points (which is true for LHV theories but not for quantum theory), then \mathcal{T} will be a convex polytope and so can be described as the intersection of the half-spaces defined by a set of “facet-defining” linear inequalities. Facet-defining inequalities are defined in terms of affinely independent points, where a set \mathcal{S} of K vectors \vec{p}_i , $\mathcal{S} = \{\vec{p}_0, \vec{p}_1, \dots, \vec{p}_{(K-1)}\}$ is *affinely independent* if for every $\vec{p}_k \in \mathcal{S}$, the $(K-1)$ vectors in the set $\{\vec{p}_i - \vec{p}_k | \vec{p}_i \neq \vec{p}_k\}$ are linearly independent [34].

Definition 2. A linear inequality is **facet-defining** for a convex polytope in \mathbb{R}^D when at least D affinely independent extreme points saturate the inequality (i.e. satisfy the equality of the linear inequality).

Having constructed a framework for stochastic maps (and equivalently, conditional probabilities) we can immediately apply this to the correlators studied in this paper. Recall that we will focus our attention on correlators $p(k|\mathbf{s})$ representing the probability distribution on $k = \sum_{j=1}^n m_j$, the sum modulo d of the outputs. The correlators have the form of a conditional probability and thus a stochastic map, and therefore are characterized by vectors in the space $\mathcal{P}_{n,c,d}$. The space $\mathcal{P}_{n,c,d}$ represents the set of all correlators possible in principle, but for certain theories, not all correlators will be permitted. The focus of our study is then the regions of correlators which are accessible in LHV theories and in quantum theories. The region of LHV correlators is itself a polytope, and the facet-defining inequalities represent the facet-defining Bell inequalities [10–12, 36]. This is the geometrical picture of Bell inequalities that we will consider in the next subsection. We now study and characterize the region of correlators in LHV theories.

B. Local Hidden Variable Theories

LHV theories are theories in which the measurement outcomes at each site can depend on the input setting of the measurement. The only other parameter the outcome depends on is an objective “hidden variable” $\lambda \in \Lambda$ which we assume is shared by all parties, and where Λ is the (often continuous) set which defines the possible values for this variable. We do not assume that λ is deterministic, but allow it to satisfy a probability distribution $p(\lambda)d\lambda$ over the space Λ such that $\int_{\Lambda} p(\lambda)d\lambda = 1$. Without loss of generality [36], we assume that apart from the probability distribution over λ the internal workings of each measurement device is deterministic. This, via a standard argument, results in conditional probabilities for the set of outputs \mathbf{m} of the following form,

$$p(\mathbf{m}|\mathbf{s}) = \int_{\Lambda} d\lambda p(\lambda) \prod_{j=1}^n p(m_j | s_j, \lambda). \quad (7)$$

Since λ is the only non-deterministic element to consider, we can study the LHV correlations as the convex combination of the set of possible deterministic maps. We now prove that the map from measurement settings to correlators in any LHV theory is a probabilistic combination of n -partite linear functions.

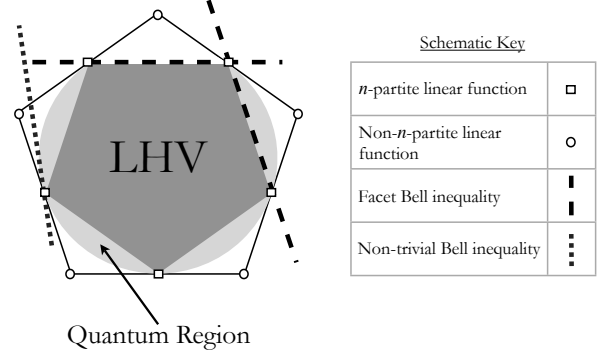


FIG. 2: Schematic for correlators in LHV theories versus more general theories (including quantum mechanical correlators). The shapes represent the convex polytopes of general correlators and LHV correlators, i.e. the space \mathcal{L} . The points on each shape are the deterministic maps corresponding to the particular functions listed. We have shown the region of quantum correlators lying between the convex polytope of n -partite linear functions and the bigger convex polytope of all possible (including non- n -partite linear) maps. The facet Bell inequalities define the region \mathcal{L} of LHV correlators whereas the non-trivial Bell inequalities (as mentioned in section III) just bound the region. It is preferable but very hard [12] to find facet Bell inequalities rather than just non-trivial ones.

Theorem 1. In an n -party generalized Bell experiment, with c settings and d outputs for each measurement, the region \mathcal{L} of correlators accessible in a LHV theory is the convex hull of the set of n -partite linear functions.

Proof. The space \mathcal{L} of correlators is a convex polytope, we only need to find the extreme points of the polytope, or deterministic correlators, and take their convex hull [36]. This means we just need to find the deterministic functions possible with LHV correlators. These deterministic functions correspond to adding the measurement outcomes of space-separated sites, which, for a LHV theory, is deterministic if and only if the maps from the measurement setting s_j to the measurement outcome m_j at each site is deterministic, i.e., of the form $g_j(s_j)$ in Definition 1 for all j . Adding such functions modulo d gives a n -partite linear function. \square

This theorem generalizes the theorem valid for the (n,2,2) case presented by two of us in [8]. It shows that \mathcal{L} is compactly characterized by considering it in terms of the computational expressiveness of LHV correlators (see Fig. 2 for a schematic of the consequences of Theorem 1). In other words, if one considers the Bell experiment as a computation, LHV theories allow the computation of n -partite linear functions and nothing more. From this perspective, LHV theories have limited computational power, and thus a theory that allows correlators outside this region can be perceived as having a *computational advantage* in a Bell experiment.

As mentioned above, there is a dual description of a convex polytope in terms of linear inequalities. The convex polytope

of LHV correlators \mathcal{L} is the intersection of the half-spaces described by the set of facet-defining Bell inequalities (or facet Bell inequalities for short). In Fig. 2 we illustrate the different types of Bell inequalities which one can construct, and the terminology to classify them. Bell inequalities that are equivalent to normalization or positivity conditions (i.e. the boundaries of \mathcal{P}) are referred to as trivial Bell inequalities. There remain Bell inequalities that are neither necessarily facet-defining or trivial, which we call *non-trivial* Bell inequalities; they are non-trivial because they indicate a separation between the polytopes \mathcal{P} and \mathcal{L} by bounding the latter.

If a correlator is outside \mathcal{L} then it must *violate* one of the facet Bell inequalities. For example, it is well-known that some quantum correlators violate a Bell inequality [1]. Also if we observe a violation of a non-trivial inequality then we know that it is outside of \mathcal{L} . To recapitulate, for either non-trivial or facet Bell inequalities, *we associate a violation with a computational advantage* and in Section III we consider Bell inequalities from this computational perspective.

Before investigating Bell inequalities, it is worth considering the relation between $p(k|\mathbf{s})$ and $p(\mathbf{m}|\mathbf{s})$. Since the latter is more general than the former, we are excluding interesting phenomena by only considering correlators. Examples of interesting phenomena in the full $p(\mathbf{m}|\mathbf{s})$ setting include the I_{3322} inequality [18] which has interesting implications with regards to quantum correlations; another example is the Guess Your Neighbor's Input [23] non-local game which results in some facet Bell inequalities that are not violated by any quantum correlators. In spite of these examples, we argue in the next subsection that the correlators $p(k|\mathbf{s})$ capture many important phenomena of full non-signaling probability distributions. In particular, we show that particular deterministic correlators can be uniquely associated with a single non-signaling distribution, that being a many-body generalization of the PR non-local box [21].

C. Non-signaling theories

In this section, we shall consider general non-signaling theories. These are theories in which there are no constraints on correlations other than the no-signaling conditions. First, note that given any deterministic correlator $p(k = f(\mathbf{s})|\mathbf{s})$ where $k = \sum_j m_j$, there exists a non-signaling distribution such that $p(m_j|s_j) = \frac{1}{d}$ and the non-signaling distribution results in the correlator. This is an immediate consequence of Theorem 2, proven below, but can be understood intuitively since every possible value of m_j is consistent with one of more output strings \mathbf{m} where $\sum_j m_j = f(\mathbf{s})$. One can then take a suitable mixture of these strings such that every output digit is maximally uncertain. Since a maximally uncertain output carries no information about \mathbf{s} , the no-signaling condition is automatically satisfied.

We denote the space of probability distributions $p(\mathbf{m}|\mathbf{s})$ that satisfy the no-signaling conditions by $\mathcal{NS}_{n,c,d}$ (the indices will normally be omitted for clarity). This space is a convex polytope as it is the intersection of the half-spaces defined by the linear inequalities being the normalization, positivity and

non-signaling inequalities (see [22] for details). Finding the vertices of this polytope is the dual problem of finding the facet-defining linear inequalities corresponding to some vertices (e.g. finding the facet Bell inequalities) which is an NP-hard problem. We now discuss some of the vertices of \mathcal{NS} without resorting to this vertex enumeration problem in the following discussion.

A famous example of a non-signaling distribution which cannot be simulated by LHV or quantum correlators for two spatially separated systems is the PR non-local box. The PR box is defined as a black box that results in measurement outcomes according to the probability distribution,

$$p(m_1, m_2 | s_1, s_2) = \begin{cases} \frac{1}{2} & \text{if } m_1 + m_2 = s_1 s_2, \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

The PR box is described by the correlator $p(k = s_1 s_2 | \mathbf{s}) = 1$. In other words, the PR box is associated with a correlator in $\mathcal{P}_{2,2,2}$. This correlator is a vertex of $\mathcal{P}_{2,2,2}$ corresponding to the deterministic property that the parity of outputs m_1 and m_2 is equal to the product of the inputs s_1 and s_2 . Furthermore, the PR box is the only distribution in $\mathcal{NS}_{2,2,2}$ which is compatible with the correlator $p(k = s_1 s_2 | \mathbf{s}) = 1$. In other words the vertex correlator $p(k = s_1 s_2 | \mathbf{s}) = 1$ in \mathcal{P} *uniquely defines* a full probability distribution in \mathcal{NS} .

In this section, we will show that this property is shared by many vertices in \mathcal{P} . These vertices are only compatible with a single probability distribution in \mathcal{NS} . In particular, for $n = 2$, every correlator in \mathcal{P} corresponding to a non- n -partite linear function is only compatible with one probability distribution in \mathcal{NS} .

We emphasize that the full set of vertices of \mathcal{NS} for arbitrary c , d and n has never been characterized, and its numerical computation is an NP hard problem (the vertex enumeration problem) [22]. For example, in the $(3, 2, 2)$ setting, the space \mathcal{NS} has been shown to have a complicated structure [37]. Therefore, we are able to capture a large amount of structure without resorting to optimization.

Before we define this structure we need to introduce a new class of functions $f(\mathbf{s}) : \mathbb{Z}_c^n \rightarrow \mathbb{Z}_d$ that we will only use in this subsection to find a *subset* of the vertices of \mathcal{NS} . We call these functions “bi-partite linear” as they correspond to functions for which, if the observers gather into two groups at two space-like separated sites, the measurement outcome results from an n -partite linear function for $n = 2$.

Definition 3. A **bipartition**, $\{A, B\}$ of the set of n parties is a division of the set $\{1, 2, \dots, n\}$ into two disjoint, non-empty sets A and B such that $A \cup B = \{1, 2, \dots, n\}$.

Definition 4. A function $f(\mathbf{s}) : \mathbb{Z}_c^n \rightarrow \mathbb{Z}_d$ is **bi-partite linear** if for any bipartition $\{A, B\}$, $f(\mathbf{s})$ can be written as

$$f(\mathbf{s}) = f^A(\mathbf{s}^A) + f^B(\mathbf{s}^B), \quad (9)$$

where \mathbf{s}^X is the $|X|$ -digit string with entries $s_j^X = s_{X_j}$ and $f^X(\mathbf{s}^X)$ is a function on \mathbf{s}^X .

We now show that there exists a unique non-signaling distribution corresponding to a vertex $p(k = f(\mathbf{s})|\mathbf{s})$ of \mathcal{P} if and only if $f(\mathbf{s})$ is not bi-partite linear.

Theorem 2. For every function $f(\mathbf{s})$, the corresponding vertex $p(k = f(\mathbf{s})|\mathbf{s}) = 1$ of \mathcal{P} is compatible with the non-signaling probability distribution

$$p(\mathbf{m}|\mathbf{s}) = \begin{cases} d^{1-n} & \text{if } \sum_{j=1}^n m_j = f(\mathbf{s}), \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

Furthermore, this is the only non-signaling distribution compatible with $p(k = f(\mathbf{s})|\mathbf{s}) = 1$ if and only if $f(\mathbf{s})$ is not bi-partite linear.

Proof. One can verify by inspection that this distribution is non-signaling (since all marginals are maximally uncertain) and satisfies $p(k = f(\mathbf{s})|\mathbf{s}) = 1$. It remains to be shown that the distribution in (10) is unique if and only if $f(\mathbf{s})$ is not bi-partite linear. To prove the if statement, note that for any bi-partite linear function $f(\mathbf{s})$, the non-signaling distribution

$$p(\mathbf{m}|\mathbf{s}) = \begin{cases} d^{2-|A|-|B|} = d^{2-n} & \text{if } \sum_{j \in A} m_j = f^1(\mathbf{s}^A) \\ & \text{and } \sum_{j \in B} m_j = f^2(\mathbf{s}^B), \\ 0 & \text{otherwise,} \end{cases} \quad (11)$$

is compatible with $p(k = f(\mathbf{s})|\mathbf{s}) = 1$. We prove the only if statement in Lemma 3. \square

Note that local operations, which, for the Bell experiment setting, corresponds to adding functions $g'_j(s_j)$ to the measurement outcomes at each site, do not change whether or not a function is bi-partite linear. Since every function can be written in terms of an n -partite linear part and a non- n -partite linear part, local operations can create the n -partite linear part and so without loss of generality this part can be set to be zero. Therefore it is sufficient to consider functions $f(\mathbf{s})$ such that $f(\mathbf{t}) = 0$ for all n -digit strings $\mathbf{t} \in \mathbb{Z}_c^n$ that have at most one non-zero digit.

We can separate the parties using any bipartition A and B and treat the sets of parties, A and B , as single parties. To do this, we take the sum modulo d of all outcomes on each side of the partition and set the number of settings on each side of the partition to be the product of the number of settings of the parties on that side of the partition (e.g., $c_A = \prod_{i \in A} c_i$). For a function that is not bi-partite linear, for all bipartitions into A and B , the function $f : \mathbb{Z}_{c_A} \times \mathbb{Z}_{c_B} \rightarrow \mathbb{Z}_d$ is non- n -partite linear for $n = 2$. Therefore it is sufficient to show that Theorem 2 applies for $n = 2$ and so it must apply for all bipartitions of n parties as long as a function $f(\mathbf{s})$ in (10) is not bi-partite linear. If the probability distribution in (10) applies for $n = 2$ then it must apply for all possible bipartitions of n parties thus generating the distribution (10) for all n .

Lemma 3. For every function $f(s_1, s_2) : \mathbb{Z}_{c_1} \times \mathbb{Z}_{c_2} \rightarrow \mathbb{Z}_d$ that is non- n partite linear for $n = 2$, the only non-signaling distribution compatible with the corresponding vertex $p(k = f(s_1, s_2)|\mathbf{s}) = 1$ in \mathcal{P} is

$$p(m_1, m_2 | s_1, s_2) = \begin{cases} d^{-1} & \text{if } m_1 + m_2 = f(s_1, s_2), \\ 0 & \text{otherwise.} \end{cases} \quad (12)$$

Proof. The condition $p(k = f(s_1, s_2)|s_1, s_2) = 1$ for all $\mathbf{s} = \{s_1, s_2\}$ implies that for every value of m_1 in $p(m_1, m_2 | s_1, s_2)$, there exists a unique value of $m_2 = f(\mathbf{s}) - m_1$. In computer science terminology we would call this a “unique game”. This immediately implies the equality for the following conditional distributions:

$$\begin{aligned} p(m_1 = x, m_2 = f(\mathbf{s}) - x | \mathbf{s}) &= p(m_1 = x | \mathbf{s}) \\ &= p(m_2 = f(\mathbf{s}) - x | \mathbf{s}), \end{aligned} \quad (13)$$

for all $x \in \mathbb{Z}_d$. The no-signaling condition further implies that $p(m_1 = x | \mathbf{s}) = p(m_1 = x | s_1)$ and

$$p(m_1 = x | s_1) = p(m_2 = f(\mathbf{s}) - x | s_2), \quad (14)$$

which must be satisfied for all \mathbf{s} and all x . We will show that repeated application of (14) for varying \mathbf{s} allows us to prove that all non-marginal probabilities are equal provided that $f(\mathbf{s})$ has a non- n -partite linear element. As discussed above, we can set $f(0, s_2) = f(s_1, 0) = 0$ for all $\mathbf{s} = \{s_1, s_2\}$. For such functions, repeatedly applying (14) gives

$$\begin{aligned} p(m_2 = -x | s_2) &= p(m_1 = x | 0) \\ &= p(m_2 = -x | 0) \\ &= p(m_1 = x | s_1) \\ &= p(m_2 = f(s_1, s_2) - x | s_2) \end{aligned} \quad (15)$$

for all x . Repeated iteration implies

$$p(m_2 = -x | s_2) = p(m_2 = \alpha f(s_1, s_2) - x | s_2) \quad (16)$$

for all $\alpha \in \mathbb{Z}_d$. The function $f(\mathbf{s})$ is non- n -partite linear so it must have at least one value of $\{s_1, s_2\}$ where $f(s_1, s_2)$ is non-zero. Since d is prime, $\alpha f(s_1, s_2)$ takes on all values in \mathbb{Z}_d , therefore the marginals are $p(m_2 | s_2) = d^{-1}$ for all m_2, s_2 . Applying equation (14) implies that $p(m_1, m_2 | s_1, s_2) = d^{-1}$ for all \mathbf{m} such that $m_1 + m_2 = f(\mathbf{s})$. \square

As we now show, for vertices $p(k = f(\mathbf{s})|\mathbf{s}) = 1$ of \mathcal{P} such that $f(\mathbf{s})$ is bi-partite linear, the existence of multiple non-signaling distributions implies that the vertex of \mathcal{P} corresponds to a facet of \mathcal{NS} .

Proposition 4. Every non-signaling distribution corresponding to a vertex of \mathcal{P} lies on a facet of \mathcal{NS} .

Proof. Any non-signaling distribution $p(\mathbf{m}|\mathbf{s})$ can be written as a convex combination $p(\mathbf{m}|\mathbf{s}) = \sum_E p(E) p_E(\mathbf{m}|\mathbf{s})$ of the vertices E of \mathcal{NS} , where $\sum_E p(E) = 1$. Therefore the associated correlator is

$$\begin{aligned} p(k|\mathbf{s}) &= \sum_E p(E) \delta_k^{\sum_{j=1}^n m_j} p_E(\mathbf{m}|\mathbf{s}) \\ &= \sum_E p(E) p_E(k|\mathbf{s}), \end{aligned} \quad (17)$$

where $p_E(k|\mathbf{s}) = \delta_k^{\sum_{j=1}^n m_j} p_E(\mathbf{m}|\mathbf{s})$ is the correlator resulting from each extreme point of \mathcal{NS} .

As the non-signaling distribution results in a correlator that is a vertex of \mathcal{P} , $p(E) = 0$ for all vertices of \mathcal{NS} that do not result in the same correlator. Denote by \mathcal{E} the set of vertices of \mathcal{NS} that correspond to the correlator and let \mathcal{F} be the region of convex combinations of the elements of \mathcal{E} . Any element in \mathcal{F} must correspond to the correlator by (17).

\mathcal{F} is not a facet of \mathcal{NS} if and only if there exists a convex combination of the vertices not in \mathcal{E} that intersects \mathcal{F} . However, if there exists such a convex combination, then this would give a convex combination of correlators in \mathcal{P} that is equal to a vertex of \mathcal{P} , which is a contradiction. \square

The uniqueness of the probability distributions in Theorem 2 implies that each vertex $p(k = f(\mathbf{s})|\mathbf{s}) = 1$ of \mathcal{P} such that $f(\mathbf{s})$ is not bi-partite linear corresponds to a unique vertex of \mathcal{NS} (i.e., the “facet” in Proposition 4 collapses to a single point).

Proposition 5. *Every probability distribution of the form (10) is a vertex of \mathcal{NS} if and only if $f(\mathbf{s})$ is not a bi-partite linear function.*

Proof. As there is a unique non-signaling distribution corresponding to the correlator $p(f(\mathbf{s})|\mathbf{s}) = 1$, there is at most one E such that $p(E) \neq 0$ in 17. \square

The structure of the high dimensional space of all possible non-signaling distributions can be partly revealed by considering the space of correlators, which are of relatively low dimension. Note that the difficulty in ascertaining the remaining structure of \mathcal{NS} comes from the difficulty of finding the set of vertices that generate vertices of \mathcal{P} that correspond to bi-partite linear functions and determining whether this gives a complete list of vertices. Furthermore, much of the structure of LHV theories in \mathcal{NS} is reduced to a small number of vertices of \mathcal{P} . In this way, the space of correlators efficiently encapsulates vital information about all non-signaling probability distributions.

There is a connection between the notion of “true n -party non-locality” [22, 38] and the vertices of \mathcal{NS} defined by Theorem 2. The bi-partite linear functions corresponding to vertices in \mathcal{P} only reveal “non-locality” (i.e. correlations not resulting from LHV theories) between a subset of all parties, but no non-locality across at least one bipartition. However, functions that are not bi-partite linear result in correlations that are non-local across all partitions.

D. Summary

In this section we have constructed a framework for Bell correlators and showed that these correlators capture important properties of the space of non-signaling probability distributions. These correlators also capture the properties of LHV theories in a natural information theoretic description, i.e. that of computational expressiveness. This work generalizes previous work of [7, 8] and shows where the structure in these previously studied settings emerges from.

Given this framework, we want to look at the correlators that are achievable with quantum theory. It is well-known

n	c	d	# Vertices	# Facet Bell inequalities
2	2	2	8	16
2	2	3	27	66
2	2	5	125	1020
3	2	2	16	256
3	2	3	81	125,412
2	3	2	32	90

TABLE I: A table of number of facet Bell inequalities for each setting (n, c, d) and the number of vertices for the LHV polytope.

that there exists quantum mechanical correlators that cannot be described by correlators resulting from an LHV theory, and this is demonstrated by the violation of a Bell inequality [1]. Therefore we now devote our attention to the Bell inequalities.

III. BELL INEQUALITIES

Bell inequalities allow us to characterize the polytope of LHV correlators. The normalization and positivity inequalities defined in Section II B define part of \mathcal{L} , the LHV polytope. However, these inequalities are also satisfied by all possible correlators $p(k|\mathbf{s})$ and define the polytope \mathcal{P} of all correlators. These inequalities then trivially bound \mathcal{L} as we cannot show a separation between \mathcal{P} and \mathcal{L} . We want to demonstrate such a separation and non-trivially bound \mathcal{P} , which we can do by constructing Bell inequalities of the form

$$\sum_{\mathbf{s}, k} \omega_{k,\mathbf{s}} p(k|\mathbf{s}) \leq \gamma_{\mathcal{L}}, \quad (18)$$

for all $p(k|\mathbf{s}) \in \mathcal{L}$, where $\omega_{k,\mathbf{s}}$ and $\gamma_{\mathcal{L}}$ are real numbers. We call $\gamma_{\mathcal{P}}$ the maximum value of the sum on the left-hand-side over correlators in \mathcal{P} . For the Bell inequality to be non-trivial, we require $\gamma_{\mathcal{L}} < \gamma_{\mathcal{P}}$.

If a correlator exists outside the LHV polytope, then it violates at least one facet Bell inequality. Finding facet Bell inequalities involves an optimization of the parameters $\omega_{k,\mathbf{s}}$ in (18) and in general is an NP-hard problem [12]. The problem is a facet enumeration problem and there are algorithmic packages that can perform this optimization such as Polymake [39]. These optimization algorithms are often used to characterize the non-signaling [22] and LHV polytopes [18] in full probability distribution settings. We used Polymake to find the facet Bell inequalities for a number of small (n, c, d) settings. The number of facet Bell inequalities for each setting is listed in Tab. I, which shows that the number of inequalities grows rapidly in the size of the problem.

For the case of $(n, 2, 2)$, the LHV polytope is a hyperoctahedron [10, 11]. In general though, there is no obvious geometrical structure or connection between the numbers of vertices and the numbers of facet inequalities. It is quite apparent that finding the facet Bell inequalities is no easy task.

In this section we will use the computational insight gained in Section II to construct Bell inequalities. We first find a

simple way to generate non-trivial Bell inequalities that necessarily bound the LHV polytope. As we shall discuss further in Section IV, these non-trivial inequalities also have a nice cross-over with the structures in so-called *non-local games*. We then describe some of the facet Bell inequalities and show that they too have a computational nature.

Finally, in the last part of this section, we discuss the quantum violation of facet Bell inequalities. We show that there is a violation for all facet Bell inequalities for $n = 2$ in Tab. I. We comment on the quantum states that *maximally* violate these Bell inequalities and find that in few instances is it a maximally entangled state. This supports the view that entanglement and a violation of a Bell inequality are not synonymous concepts.

A. Non-trivial Bell inequalities

One might ask what is the point of looking for Bell inequalities that are not facet-defining? Firstly, bounding the set of LHV correlators still creates a non-trivial subregion of a correlator space. Secondly, we can still demonstrate non-classical behavior without the computational difficulty that the facet-defining condition creates. Finally, despite not being facet-defining, non-trivial Bell inequalities can have a role in particular applications such as in non-local games or Measurement-based Quantum Computing (MBQC) [8], as non-trivial Bell inequalities might capture computational power. We will elaborate on this final point in Section IV.

In this subsection, we will discuss non-trivial Bell inequalities as being inequalities in the space \mathcal{P}' of all possible correlators $p(k|\mathbf{s})$, including those for $k = 0$. In this space, the normalisation conditions become $\sum_k p(k|\mathbf{s}) = 1$ for all \mathbf{s} , which defines a bounded hyperplane in \mathbb{R}^{dc^n} . The space of correlators \mathcal{P}' is a convex polytope defined by the positivity conditions within this hyperplane. We regard the region of LHV correlators, \mathcal{L} , as a sub-region in both \mathcal{P} and \mathcal{P}' as one can always go between the two spaces by disregarding $p(0|\mathbf{s})$ and applying the normalisation inequalities. In the rest of the paper, including the discussion on facet Bell inequalities, we always discuss correlators in the space \mathcal{P} .

Given the motivation for finding non-trivial Bell inequalities we now present a simple way of generating non-trivial Bell inequalities. We begin by considering the CHSH inequality [9] as an example. The CHSH inequality can be written as

$$\sum_{s_1, s_2} \sum_{k=0}^1 \delta_{s_1 s_2}^k p(k|s_1, s_2) \leq 3, \quad (19)$$

which is (18) with $\gamma_{\mathcal{L}} = 3$ and $\omega_{k,s} = \delta_{s_1 s_2}^k$. The CHSH inequality is also a facet Bell inequality for this setting.

By convexity, it is only necessary to consider the vertices of the LHV polytope to obtain the bound $\gamma_{\mathcal{L}} = 3$. In this case, the vertices correspond to the linear boolean functions of s_1 and s_2 [7]. For these vertices the correlators are then $p(k|s_1, s_2) = \delta_{g(s_1, s_2)}^k$ where $g(s_1, s_2)$ are the linear Boolean functions on $\{s_1, s_2\}$. The sum in (19) is then

$\sum_{s_1, s_2} \delta_{s_1 s_2}^{g(s_1, s_2)} \leq 3$ as the functions $g(s_1, s_2)$ overlaps with the function $s_1 s_2$ for at most 3 values of s . An example of a linear function achieving this overlap is $g(s_1, s_2) = 0$. On the other hand, if a correlator $p(k = f(\mathbf{s})|s_1, s_2)$ achieves the map $f(\mathbf{s}) = s_1 s_2$ deterministically, then it achieves a value of $\gamma_{\mathcal{P}} = 4$ for the sum in (19). This inequality highlights the inability for LHV correlators to evaluate non-linear Boolean functions deterministically.

For all settings (n, c, d) , we can generalize the above concepts from the Boolean algebra for $(2, 2, 2)$ to our more general framework of maps $f : \mathbb{Z}_c^n \rightarrow \mathbb{Z}_d$. We construct non-trivial Bell inequalities in the following way:

$$\sum_{\mathbf{s}} \sum_{k=0}^{d-1} \delta_{f(\mathbf{s})}^k p(k|\mathbf{s}) \leq \sup_{g(\mathbf{s})} \sum_{\mathbf{s}} \delta_{g(\mathbf{s})}^{f(\mathbf{s})}, \quad (20)$$

where the supremum is over the set of n -partite linear functions and $f(\mathbf{s})$ is a non- n -partite function of \mathbf{s} . Therefore there are $N = d^{c^n} - d^{n(c-1)+1}$ non-trivial Bell inequalities of the form (20).

We can prove the inequality in (20) using the same arguments as with the CHSH inequality. By convexity we only need to consider the vertices of \mathcal{L} which correspond to the n -partite linear functions $g(\mathbf{s})$. The sum in (20) is then $\sum_{\mathbf{s}} \sum_{k=0}^{d-1} \delta_{f(\mathbf{s})}^k \delta_{g(\mathbf{s})}^k$. We then take the supremum over all n -partite functions $g(\mathbf{s})$ to obtain the right-hand-side of (20).

Whilst these inequalities might seem contrived, they include some interesting Bell inequalities. For example, the Svetlichny inequality [38] for three parties takes the form of (20), i.e.,

$$\sum_{s_1, s_2, s_3} \delta_{s_1(s_2+s_3)+s_2 s_3}^k p(k|s_1, s_2, s_3) \leq 6, \quad (21)$$

which is (20) for $f(\mathbf{s}) = s_1(s_2+s_3)+s_2 s_3$. The Svetlichny inequality is not a facet Bell inequality for the LHV polytope \mathcal{L} in the $(3, 2, 2)$ setting [10] but it still captures very interesting phenomena. In particular, the Svetlichny inequality captures correlations that are consistent with LHV theories if one averages over the measurement settings and outcomes for any one of the parties.

The non-trivial Bell inequalities (20) can be seen to truly capture the inability of LHV theories to evaluate non- n -partite linear functions via correlators. A way of emphasizing this is to consider these inequalities as a “non-local game” [5]. That is, given an input \mathbf{s} to n parties that do not communicate, what is the maximum probability of the parties producing outcomes m such that $\sum_{j=1}^n m_j = f(\mathbf{s})$ for some function $f(\mathbf{s})$ using different strategies (or resources such as shared randomness or entanglement). A referee chooses the input and then obtains all outputs m before computing $\sum_{j=1}^n m_j$.

If the referee’s choice of input is uniformly random, then the inequality in (20) captures the mean success probability of an LHV strategy performing the function $f(\mathbf{s})$; one just needs to divide both sides by c^n to obtain a proper probability. If one imagines that the choice of input \mathbf{s} is not chosen at random but instead with some probability $p(\mathbf{s}) \geq 0$ such that $\sum_{\mathbf{s}} p(\mathbf{s}) = 1$

then we can modify (20) to be:

$$\sum_{\mathbf{s}} p(\mathbf{s}) \sum_{k=0}^{d-1} \delta_{f(\mathbf{s})}^k p(k|\mathbf{s}) \leq \sup_{g(\mathbf{s})} \sum_{\mathbf{s}} p(\mathbf{s}) \delta_{g(\mathbf{s})}^f, \quad (22)$$

to get the mean success probability of performing the function $f(\mathbf{s})$ with LHV correlators given a distribution $p(\mathbf{s})$ on inputs. We now present the following proposition that shows an infinite class of non-trivial Bell inequalities can be easily generated.

Proposition 6. *Any Bell inequality of the form (22) is non-trivial whenever $p(\mathbf{s}) \neq 0$ for all \mathbf{s} .*

Proof. For any n -partite linear function,

$$\sum_{\mathbf{s}} p(\mathbf{s}) \delta_{g(\mathbf{s})}^f < \sum_{\mathbf{s}} p(\mathbf{s}) = 1 \quad (23)$$

as there must be a value of \mathbf{s} such that $\delta_{g(\mathbf{s})}^f = 0$. Therefore the right-hand-side of (22) is strictly less than 1 for all correlators in \mathcal{L} . However, for the correlator $p(k = f(\mathbf{s})|\mathbf{s}) = 1$ in \mathcal{P} , the left-hand side of (22) is exactly 1. \square

The beauty of these non-trivial Bell inequalities is that they are easily generated as one can easily describe the n -partite linear functions for any setting (n, c, d) , yet they still bound the region of LHV correlators. They also explicitly capture the computational aspect of a Bell inequality experiment: the value of the left hand side of (22) gives the mean success probability of performing a particular function.

We now consider the facet Bell inequalities that completely characterize the LHV polytope for small values of (n, c, d) . Whilst these latter inequalities do not have the immediacy of the non-trivial Bell inequalities and are difficult to calculate, they are still the optimal tool for demonstrating non-classical correlations.

B. Facet Bell inequalities

The facet Bell inequalities satisfy the facet-defining condition for LHV correlators and so if a correlator is outside the region of LHV correlators it must necessarily violate one of the facet Bell inequalities. In Tab. I we listed the number of facet Bell inequalities for a few settings that could be computed using the software Polymake. Included in the number of facet Bell inequalities are the c^n normalization and $(d-1)c^n$ positivity inequalities that define the polytope \mathcal{P} . Despite these dc^n inequalities, there are still a significant number of inequalities remaining. Also note we are again discussing correlators in the space \mathcal{P} of all correlators except $p(0|\mathbf{s})$ and \mathcal{L} is the sub-region of LHV correlators in this space.

We can reduce the number of facet Bell inequalities by considering relabelings of the measurement choices, outcomes and parties, which we refer to as symmetries. If a Bell inequality can be changed into another inequality by such an operation, we refer to them as equivalent, or as members of the same symmetry class. Explicitly, the symmetries are

n	c	d	# Symmetry classes
2	2	2	1
2	2	3	1
2	2	5	4
3	2	2	4
3	2	3	62
2	3	2	1

TABLE II: The number of symmetry classes for each setting (n, c, d) . We have excluded the symmetry class of all positivity and normalisation inequalities from this number.

1. permutations of parties - $\{s_i, s_j, \dots, s_n\} \rightarrow \{s_{i'}, s_{j'}, \dots, s_{n'}\}$ where $k' = \sigma(k)$ is an element of the permutation group of order n
2. relabeling of measurement settings - $s_j \rightarrow s_j + \alpha_j$ for some $\alpha_j \in \mathbb{Z}_c$
3. relabeling of measurement outcomes - $m_j \rightarrow m_j + \beta(s_j, j)$ where $\beta(s_j, j) \in \mathbb{Z}_d$.

The n -partite linear functions are closed under all of these operations. Using the facet-defining condition, the vertices of \mathcal{L} that saturate a facet Bell inequality must be equivalent to another set of vertices in \mathcal{L} by these operations that saturate another facet Bell inequality. Hence we can group together the facet Bell inequalities into these symmetry classes.

There are $n!$ permutations of n parties and c^n ways of relabeling measurement settings. Since for each input s_j we add a value $\beta(s_j)$, for each input \mathbf{s} , $\beta(\mathbf{s}) = \sum_{j=1}^n \beta(s_j)$ is added to $\sum_{j=1}^n m_j$. There will be at most d^{cn} values of $\beta(\mathbf{s})$.

We constructed a search algorithm to find all the symmetry classes of facet Bell inequalities for each setting (n, c, d) . In Tab. II we list the number of symmetry classes for each of the settings in Tab. I, which is dramatically fewer than the total number of facet Bell inequalities. One of the symmetry classes for each setting is the class of normalization and positivity conditions, leaving all the facet Bell inequalities that can be violated by LHV correlators outside of \mathcal{L} . We have therefore omitted the class of these positivity and normalization inequalities from Tab. II in order to leave the inequalities that can be violated.

We now look at a few particular families of facet Bell inequalities for the bipartite and tripartite setting based on the CGLMP inequality. We briefly discuss the $(2, 3, 2)$ setting and show how the CHSH inequality is essentially the only relevant facet inequality for this setting.

Bipartite Two-Setting Facet Bell Inequalities - A symmetry class of note for $n = c = 2$ is the class of facet Bell inequalities that are equivalent to the CGLMP inequality [14]. We can rewrite this inequality after symmetry transformations and

mapping it into our normalized probability space \mathcal{P} to obtain,

$$\begin{aligned} \mathcal{C}_{CGLMP} = & d \times p\left(\sum_{j=1}^2 m_j = 1|0,0\right) - \\ & \sum_{\mathbf{s}} (-1)^{s_1+s_2} p\left(\sum_{j=1}^2 m_j = 1|\mathbf{s}\right) + \\ & \sum_{\mathbf{s}} (-1)^{s_1+s_2} \sum_{k=2}^{d-1} (d-k-1) p\left(\sum_{j=1}^2 m_j = k|\mathbf{s}\right) \leq d. \end{aligned} \quad (24)$$

This inequality is also equivalent to the CHSH inequality for $d = 2$ which we now write in the space \mathcal{P} for clarity:

$$\mathcal{C}_{d=2} = \sum_{\mathbf{s}} (-1)^{s_1+s_2} p\left(\sum_{j=1}^2 m_j = 1|\mathbf{s}\right) \leq 2. \quad (25)$$

For all possible correlators in \mathcal{P} , the maximal value of the left-hand-side of the CGLMP inequality is $2d - 1$, thus violating it. In fact, for all d , this maximal violation of the CGLMP is obtained by a vertex of \mathcal{P} corresponding to the map $f(\mathbf{s}) = s_1 s_2 + 1$, i.e. the correlator $p(k|\mathbf{s}) = \delta_{s_1 s_2 + 1}^k$. Masanes has shown that the only non-trivial, facet Bell inequalities are those that are equivalent to the CGLMP for the $(2, 2, 3)$ setting [40]; this is confirmed by Tab. II.

For the $(2, 2, 2)$ setting there are $2^4 - 2^3 = 8$ non- n -partite linear functions and also 8 inequalities in the symmetry class of the CHSH inequality. This is no coincidence as every Bell inequality in the symmetry class is maximally violated by a vertex of \mathcal{P} corresponding to a non- n -partite linear function. This also occurs for the $(2, 2, 3)$ setting, there are $3^4 - 3^3 = 54$ non- n -partite linear functions and 54 inequalities in the symmetry class of the CGLMP inequality. It can also be checked that every inequality in this symmetry class is violated by a different non- n -partite linear function.

The correspondence between non- n -partite linear functions and facet Bell inequality echoes the non-trivial Bell inequalities described by the inequalities of the form (20). There is a computational aspect to each inequality as a violation indicates that a correlator is in some sense closer to the non- n -partite linear functions. However, the correspondence breaks down for the $(2, 2, 5)$ setting where there are $5^4 - 5^3 = 500$ non- n -partite linear functions but 1000 non-trivial, facet Bell inequalities in total. In this instance, each vertex of \mathcal{P} maximally violates two facet Bell inequalities (each belonging to different symmetry classes). Therefore there are 1000 facet Bell inequalities (see Appendix A for more details).

Tripartite Facet Bell Inequalities - We have given an indication that facet Bell inequalities for $n = c = 2$ have a computational interpretation. Every facet Bell inequality we have found is maximally violated uniquely by a vertex of \mathcal{P} . In this sense the violation of a facet Bell inequalities can quantify how computationally powerful a theory is. For situations with $n > 2$, this becomes more complicated even for $n = 3$ and $c = d = 2$. The Mermin inequality [41] for example is in

the symmetry class of the inequality

$$\sum_{\mathbf{s}} \delta_{s_2}^{s_1} (-1)^{s_1+s_3} p\left(\sum_{j=1}^3 m_j = 1|\mathbf{s}\right) \leq 2. \quad (26)$$

Two vertices of \mathcal{P} that are not vertices of \mathcal{L} maximally violate this inequality. The two maps that correspond to these vertices are $f(\mathbf{s}) = s_1 s_2 s_3 + 1$ and $f(\mathbf{s}) = s_1 s_3 + 1$. Therefore there is no longer a one-to-one correspondence between facet inequalities and the vertices of \mathcal{P} that violate the inequality.

The CHSH inequality expressed in terms of the expectation value of measurements (see next subsection) for two parties can be used to build facet Bell inequalities for more parties [10] such as the above inequality (26). Analogously, we define CGLMP for three parties using the two party inequality. We have three parties but now we only consider non-zero terms in a Bell inequality when the third party's measurement setting $s_3 = 0$. For LHV correlators $p(k|s_1, s_2, 0)$ the n -partite linear functions that can be achieved are $f(\mathbf{s}) = [\alpha_1 s_1 + \alpha_2 s_2 + \alpha_3]_3$ with $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{Z}_d$: the n -partite linear functions on two variables s_1 and s_2 . Since the CGLMP inequality is facet-defining for the region of LHV correlators for two parties, or variables s_1 and s_2 , it is facet-defining for this space of the $n = 3$ correlators for $s_3 = 0$. Then we can write the tripartite CGLMP inequality as

$$\begin{aligned} \mathcal{C}'_{CGLMP} = & d \times p\left(\sum_{j=1}^3 m_j = 1|0,0,0\right) - \\ & \sum_{\mathbf{s}} (-1)^{s_1+s_2} p\left(\sum_{j=1}^3 m_j = 1|s_1, s_2, 0\right) + \\ & \sum_{\mathbf{s}} (-1)^{s_1+s_2} \sum_{k=2}^{d-1} (d-k-1) p\left(\sum_{j=1}^3 m_j = k|s_1, s_2, 0\right) \leq d. \end{aligned} \quad (27)$$

For the case of $(3, 2, 3)$, this tripartite CGLMP inequality is facet-defining and forms a symmetry class with 324 inequalities. This is also true for $d = 2$ where \mathcal{C}'_{CGLMP} results from the CHSH inequality and forms one of 4 non-trivial symmetry classes. Also in this case there is another way of substituting the CHSH inequality to obtain the class containing the inequality (26). This class also generalizes to the $d = 3$ setting so that the following inequality is against the space where $s_1 = s_2$ in a similar fashion to \mathcal{C}'_{CGLMP} :

$$\begin{aligned} \mathcal{C}''_{CGLMP} = & d \times p\left(\sum_{j=1}^3 m_j = 1|0,0,0\right) - \\ & \sum_{\mathbf{s}} \delta_{s_2}^{s_1} (-1)^{s_1+s_3} p\left(\sum_{j=1}^3 m_j = 1|\mathbf{s}\right) + \\ & \sum_{\mathbf{s}} \delta_{s_2}^{s_1} (-1)^{s_1+s_3} \sum_{k=2}^{d-1} (d-k-1) p\left(\sum_{j=1}^3 m_j = k|\mathbf{s}\right) \leq d. \end{aligned} \quad (28)$$

This inequality also forms a symmetry class with 324 inequalities. There are 60 other symmetry classes for $(3, 2, 3)$ (excluding the class of positivity and normalization inequalities). Inequalities from each of these classes can be found in the supplementary material [42].

Bipartite Three-Setting, Two-Outcome Facet Inequalities - Finally, as can be seen from Tab. II for the $(2, 3, 2)$ setting there is only one symmetry class of non-trivial facet Bell inequalities. The Bell inequality generating this symmetry class is a generalization of the CHSH inequality:

$$\mathcal{C}_{c=3} = \sum_{\mathbf{s}} (-1)^{s_1 s_2} \prod_{j=1}^2 (\delta_0^{s_j} + \delta_1^{s_j}) p\left(\sum_{j=1}^2 m_j = 1 | \mathbf{s}\right) \leq 2. \quad (29)$$

This is exactly the same as the CHSH inequality if instead of $c = 3$ we had $c = 2$. In fact for either c or d equal to 4 and $n = 2$, inequalities constructed from the CHSH inequality capture a lot of the structure of the LHV polytope.

In this subsection we have given an indication of the richness of the structure of the LHV polytope for some simple settings. There is also a computational element to some of these Bell inequalities, such as the CGLMP inequality which is maximally violated by a vertex of \mathcal{P} , in that greater violations enable correlators to come closer to evaluating a non- n -partite linear function deterministically. Quantum correlators are known to provide violations of Bell inequalities and so in the next section we discuss quantum correlators.

C. Quantum correlators

Quantum correlations can violate all manner of Bell inequalities if the correlations are generated from measurements on an entangled state. However, the connection between a violation of a facet Bell inequality and entanglement is not completely clear. Recently the two concepts have become very distinct and nowhere is this best demonstrated by the CGLMP inequality where the greatest violation of the CGLMP inequality is *not* achieved with a bipartite maximally entangled state. The region $\mathcal{Q} \subset \mathcal{P}$ of quantum correlators is convex, but it is not a polytope since there are an infinite number extreme points [12]. It is not clear in general how to define all of the extreme points of \mathcal{Q} . We now discuss this region in the context of our CGLMP-like Bell tests.

Without loss of generality, we can assume that the measurements at each site are projective measurements by Naimark's theorem [43], so that for a quantum state $\rho \in (\mathcal{H})^{\otimes n}$, the correlators are

$$p_{\mathcal{Q}}(k | \mathbf{s}) = \sum_m \delta_k^{\sum_{j=1}^n m_j} \text{tr}(\rho \bigotimes_{j=1}^n |m_j\rangle_{s_j} \langle m_j|_{s_j}). \quad (30)$$

As ρ can be expressed as a convex combination of pure states, we can, for the purposes of finding the maximum quantum violation, assume that ρ is a pure state.

A more compact way of expressing Bell inequalities in terms of the expectation values of joint measurements M_{s_j}

[15, 35], which, for the above projective measurements, are

$$\begin{aligned} \mathbb{E}(\mathbf{s}) &= \text{tr}(\rho \bigotimes_{j=1}^n M_{s_j}) \\ &= \sum_m e^{\frac{i2\pi}{d} (\sum_{j=1}^n m_j)} \text{tr}(\rho \bigotimes_{j=1}^n |m_j\rangle_{s_j} \langle m_j|_{s_j}) \\ &= \sum_{k=0}^{d-1} e^{\frac{i2\pi k}{d}} p_{\mathcal{Q}}(k | \mathbf{s}). \end{aligned} \quad (31)$$

The correlators $p_{\mathcal{Q}}(k | \mathbf{s})$ in (31) can be replaced with LHV correlators to obtain the expectation value of measurements for LHV theories.

Our construction of correlators therefore has a natural role in the construction of expectation values $\mathbb{E}(\mathbf{s})$. If we have a Bell inequality in terms of correlators such as (18), it is possible to relate it to a Bell inequality in terms of these expectation values by taking the discrete Fourier transform [35]. The discrete Fourier transform of (18) gives

$$\sum_s \sum_{\mu=0}^{d-1} \eta_{\mu, \mathbf{s}} [\mathbb{E}(\mathbf{s})]^\mu \leq \gamma_{\mathcal{L}} \quad (32)$$

with $\gamma_{\mathcal{L}}$ as defined in (18) if the complex pre-factors $\eta_{\mu, \mathbf{s}}$ are

$$\eta_{\mu, \mathbf{s}} = \frac{1}{d} \sum_{k=1}^{d-2} \omega_{k, \mathbf{s}} e^{-i2\pi \frac{\mu k}{d}}. \quad (33)$$

This construction is another motivation for considering correlators as opposed to the full distribution $p(\mathbf{m} | \mathbf{s})$.

To obtain the maximal violation of a Bell inequality by quantum correlators, we optimize (32) over pure states ρ and unitary operators M_{s_j} (which correspond to projective measurements). For the case of $(n, 2, 2)$, all facet Bell inequalities are maximally violated by the GHZ state $|GHZ\rangle = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes n} + |1\rangle^{\otimes n})$ [10]. In general, finding the maximal violation is a difficult problem, however, there are methods of providing numerical lower and upper bounds on this quantum violation (see e.g. [45] and [15]). We will now discuss and utilize these methods to find the maximal quantum violations of facet Bell inequalities for two parties.

First we discuss methods of finding a lower bound on a two-party maximal quantum violation used in [15] and [46]. The quantum state is first fixed as the d^2 -dimensional maximally entangled state $|\Psi\rangle = \frac{1}{\sqrt{d}} \sum_{j=0}^{d-1} |jj\rangle$ and we optimize over the unitaries M_{s_j} . More specifically we write the projectors as $|m_j\rangle_{s_j} \langle m_j|_{s_j} = M_{s_j} |k\rangle \langle k| M_{s_j}^\dagger$, where $\{|k\rangle | k \in \mathbb{Z}_d\}$ is the standard basis of \mathcal{H}^D . The M_{s_j} can be written as $M_{s_j} = F D_{s_j}$ where F is the d -by- d Hadamard, or Quantum Fourier Transform matrix and $D_{s_j} = \text{diag}(e^{i\phi_1(s_j)}, e^{i\phi_2(s_j)}, \dots, e^{i\phi_d(s_j)})$, a diagonal matrix with $\phi_j(s_j)$ as real phases. Therefore we optimise over these phases $\phi_j(s_j)$ to numerically maximize the quantum violation for the maximally entangled state.

Once we find the optimal phases $\phi_j(s_j)$ for each s_j , we optimize over the pure states, ρ , by finding the largest eigenvalue

n	c	d	Symmetry class	LHV bound	Quantum bound
2	2	2	$\mathcal{C}_{d=2}$	2	2.4142 [†]
2	2	3	\mathcal{C}_{CGLMP}	3	3.9149
2	2	5	\mathcal{I}_1	5	6.3145
2	2	5	\mathcal{I}_2	5	7.6290
2	2	5	\mathcal{I}_3	5	7.0314
2	2	5	\mathcal{C}_{CGLMP}	5	7.0314
2	3	2	$\mathcal{C}_{c=3}$	2	2.4142 [†]

TABLE III: We list the bipartite maximal quantum violations for particular facet Bell inequalities for c and d . The inequalities \mathcal{I}_1 , \mathcal{I}_2 and \mathcal{I}_3 can be found in the appendix A. Those violations that are achieved with the bipartite maximally entangled state of d^2 dimension are labeled with a [†].

of the operators corresponding to (32). The largest eigenvalue then corresponds to the eigenvector $|\psi\rangle$. Using this method it has been shown that the maximal violation of the CGLMP inequality is achieved with a state that is not maximally entangled [46].

One method of providing an upper bound to the violation of a Bell inequality is the use of semi-definite programming [45]. A Gram matrix is constructed from the expectation value of products (or sequences) of projectors and this matrix is positive semi-definite. The Bell inequality is then a linear objective function on this matrix subject to linear constraints on elements of the Gram matrix. Navascues et al constructed a hierarchy of semi-definite programs that converge to the set of quantum correlations and also constructed a method for testing for convergence [45].

We used YALMIP and SeDuMi [47] to implement these semi-definite algorithms to find the upper bound to quantum violations whilst finding lower bounds using previously described methods. In Tab. III, we list the maximal quantum violations of the facet Bell inequalities in symmetry classes described previously; a single value is listed as the upper and lower bounds coincide differ by at most 10^{-9} , which is consistent with numerical error. Interestingly, of the settings studied, only Bell inequalities for qubits are maximally violated by a maximally entangled bipartite state.

D. Summary

Dual to a vertex description of a convex polytope is the facet description where facets are defined by linear inequalities. We have given an insight into some of the complicated structure behind these facet Bell inequalities. Despite the complicated structure, there is a computational insight into the facet Bell inequalities for particular settings. The CGLMP inequalities are computational in nature and can be used to construct facet Bell inequalities for multipartite scenarios. We then showed that these and other bipartite facet Bell inequalities are violated by quantum correlators, which indicates that quantum mechanics offers a computational advantage relative to any LHV theory.

The non-trivial Bell inequalities that we have constructed in this section have an explicit computational meaning. We will explore these inequalities in the context of non-local games and Measurement-based Quantum Computing. More generally we consider how much of the structures developed in this paper can be applied to these information processing scenarios. Whilst a lot of ideas have been generalized from the $(n, 2, 2)$ setting, we show that LHV correlators are more powerful when $c, d > 2$.

IV. BELL TESTS, NON-LOCAL GAMES AND QUANTUM COMPUTING

Connections have been made between Bell tests and Measurement-based Quantum Computing (MBQC) [6, 8]. This connection has been explicitly explored in the setting where each party has a choice of two measurements, each with two outcomes. Furthermore, the role of post-selection in Bell tests simulating MBQC has been explored in this setting [7] leading to novel quantum phenomena. In this section we will give an overview of all of these connections and then discuss their implementation in general (n, c, d) settings. So much of the computational insight in constructing LHV correlators in the $(n, 2, 2)$ setting carries over into the (n, c, d) setting, but the structure of LHV correlators is also richer. This richness also means certain results do not generalize.

We have already discussed non-local games in the context of the non-trivial Bell inequalities in section III. There is an overlap between non-local games and MBQC if we consider the elements in both information processing scenarios. Firstly in MBQC there are a number of sites that share a particular resource (e.g. the cluster state [48] for quantum computing), and single-site measurements are made on this resource. All measurement data is processed by a classical computer and in order to achieve a universal quantum computer in current models, adaptive measurements are required; adaptivity means choices of measurements are informed by previous measurement outcomes. In the model developed by Raussendorf and Briegel, the classical computer only needs to be able to perform linear Boolean functions to achieve a universal quantum computer [48].

We now recall the model of non-local games as discussed earlier [5]. There are a number of parties who share some resource state but do not communicate with each other. These parties receive an input from a referee and send an output back. The referee processes the outputs to see if the parties successfully performed some task. MBQC can be recast as a non-local game where the parties choose their measurement based upon the input they receive and send the referee the measurement outcome. The referee in MBQC is a classical computer who processes the measurement data in order to achieve some task. Therefore the referee performs linear Boolean function computation on the measurement data described by bits [8].

MBQC has been generalized to include more than a choice of two measurements and two measurement outcomes at each site [49], thus measurement data is no longer encoded in bits.

The classical computer processing measurement data in these models just uses addition modulo d where d is the number of measurement outcomes at each site [49]. Addition modulo d on data is still an extremely limited form of computation, even compared to the n -partite linear functions discussed in this paper.

For simplicity we say that the data sent and received from the measurements sites is encoded as digits in \mathbb{Z}_d . The specific model of MBQC as a non-local game that we consider consists of three stages:

1. A digit-string $\mathbf{x} \in \mathbb{Z}_d^{|\mathbf{x}|}$ of length $|\mathbf{x}|$ is processed by a classical computer, which then sends a single digit to each site after processing;
2. At each site, this digit is used to select a measurement and the outcome is sent to the classical computer;
3. The classical computer processes this measurement data to output a single digit.

The processing power of the classical computer is limited to addition modulo d .

The input at each site $s_j = h(\mathbf{x})$ is now a result of some pre-processing leading to a function $h(\mathbf{x})$ on the digit-string \mathbf{x} where $|\mathbf{x}| \leq n$. This function is limited to addition modulo d on \mathbf{x} , i.e. $h(\mathbf{x}) = \sum_{j=1}^{|\mathbf{x}|} \alpha_j x_j$ where $\alpha_j \in \mathbb{Z}_d$. Then the processing on measurement outcomes m_j leads to the output function $\sum_{j=1}^n m_j$ without loss of generality.

The non-trivial Bell inequalities described in this paper capture the mean success probability of LHV correlators evaluating a non- n -partite linear function $f(\mathbf{s})$ for the input \mathbf{s} . We now investigate how the construction of Bell inequalities is modified by this new element of pre-processing on a digit-string \mathbf{x} . The pre-processing allows us to express the correlators as being conditioned upon \mathbf{x} , i.e., $p(k|\mathbf{x}) := p(k|\mathbf{s}(\mathbf{x}))$.

The correlators $p(k|\mathbf{x})$ are now elements of a vector \vec{p} in $\mathbb{R}^{(d-1)d^{|\mathbf{x}|}}$. We now describe the structure of the region of all possible correlators.

For the digit-string \mathbf{x} if we set $n = |\mathbf{x}|$ and $s_j = x_j$ for all j , then the possible correlators $p(k|\mathbf{x})$ live in the space \mathcal{P} for this setting. If we increase n and consider other forms of linear pre-processing, we allow more freedom to evaluate more complicated functions. As shown in [8] for the pre-processing described above, quantum correlators for $d = 2 = c$ benefit from this freedom. That is, for $d = 2 = c$ settings, all deterministic correlators $p(k|\mathbf{x})$ corresponding to functions $f(\mathbf{x}) : \mathbb{Z}_d^{|\mathbf{x}|} \rightarrow \mathbb{Z}_d$ can be achieved for some sufficiently large n .

Remarkably, LHV correlators $p(k|\mathbf{x})$ for $d = 2$ are unaffected by increasing n using the pre-processing described above [7]. The deterministic correlators, or vertices of \mathcal{P} for $p(k|\mathbf{x})$ that LHV correlators achieve are always the linear Boolean functions. In this way, the computational power of LHV correlators is not boosted by increasing n and pre-processing for $d = 2$. We now show that for general $d > 2$ this no longer holds and computational power can be boosted. We actually show something stronger: if any form of the pre-processing described above with $s_j = h(\mathbf{x})$ and $|\mathbf{x}| = n$ is al-

lowed, then the computational power of LHV correlators will be boosted.

Proposition 7. *For arbitrary addition modulo d pre-processing on \mathbf{x} leading to measurement settings $s_j = h(\mathbf{x})$, the space of LHV correlators $p(k|\mathbf{x})$ is not confined to the convex hull of n -partite linear functions on \mathbf{x} for $d > 2$.*

Proof. As before, we consider the deterministic maps $p(k|\mathbf{s})$ and then take their convex hull. However, we now consider the effect of pre-processing on \mathbf{x} and make the assumption that $d > 2$. When each party receives the input s_j generated by the pre-processing, the input is $s_j = h(\mathbf{x})$ where $h(\mathbf{x}) = \sum_{j=1}^{|\mathbf{x}|} \alpha_j x_j$ with $\alpha_j \in \mathbb{Z}_d$. We now map into the space of all correlators $p(k|\mathbf{x})$ under the influence of all possible pre-processing of the form $h(\mathbf{x})$.

For LHV correlators, we need to consider all single site maps $\mathbb{Z}_d \rightarrow \mathbb{Z}_d$ which can be written as a polynomial over \mathbb{Z}_d as:

$$m_j = \sum_{y=1}^{(d-1)} \sum_{z=0}^{(d-1)} \epsilon_{y,z} s_j^{d-(z+1)} + \gamma, \quad (34)$$

where $\epsilon_{y,z} = \beta_y (-1)^{z+1} \binom{d-1}{z} y^z$ and $\gamma, \beta_y \in \mathbb{Z}_d$. If we add in the pre-processing stage where $s_j = \sum_{j=1}^{|\mathbf{x}|} \alpha_j x_j$ then single site maps become polynomials in elements of x_j , i.e. $s_j^{d-(z+1)} = (\sum_{j=1}^{|\mathbf{x}|} \alpha_j x_j)^{d-(z+1)}$. Therefore for appropriately chosen β_y and α_j , there are now cross-terms between elements of x_j , e.g. $x_1 x_2$ etc if $d > 2$. Because of this the deterministic single-site maps with pre-processing cannot be described as n -partite linear functions on \mathbf{x} . \square

If we just have $s_j = x_j$ and have $n = |\mathbf{x}|$ parties then the LHV polytope is just the convex hull of n -partite linear functions on \mathbf{x} . Pre-processing therefore can boost the computational power of LHV correlators to go beyond this computational description. However, not all polynomials of elements x_j over \mathbb{Z}_d can be achieved by taking powers of the linear functions $\sum_{j=1}^{|\mathbf{x}|} \alpha_j x_j$. For example, the function $f(\mathbf{x}) = \prod_{j=1}^{|\mathbf{x}|} x_j^{d-1}$ cannot be achieved by this method. More generally, any function that contains a term that is a non- n -partite linear function of degree greater than $d - 1$ cannot be achieved in any LHV theory for any value of n . So whilst the LHV correlators are boosted by this pre-processing, the resulting LHV polytope does not encompass the convex hull of all functions $f(\mathbf{x})$.

Knowing that the boosted LHV polytope cannot achieve a function such as $f(\mathbf{x}) = \prod_{j=1}^{|\mathbf{x}|} x_j^{d-1}$, we can construct a non-trivial Bell inequality for this function for all $n \geq |\mathbf{x}|$. Instead of finding the upper bound for LHV correlators by taking the supremum over all n -partite linear functions, we now have to take the supremum over all possible linear pre-processing for all n maps of the form (34). This added complication highlights the uniqueness and simplicity of the case for $d = 2$, where one only needs to take the supremum over all linear Boolean functions of \mathbf{x} even with pre-processing.

In the case where the number of possible inputs at each site is different from the number of possible outputs, there are other complications. We illustrate this with an example in the $(3, 2, 3)$ setting but with a bit-string $\mathbf{x} \in \mathbb{Z}_2^2$. We have three parties with pre-processing on \mathbf{x} leading to inputs: $s_1 = x_1$, $s_2 = x_2$ and $s_3 = [x_1 + x_2]_2$. With this pre-processing one can achieve the non- n -partite linear function $f(\mathbf{x}) = x_1 x_2$ with the n -partite linear function $f(\mathbf{s}) = 2s_1 + 2s_2 + s_3$ and this pre-processing as $2x_1 + 2x_2 + [x_1 + x_2]_2 = \delta_1^{x_1 x_2}$. Therefore even with $c = 2$ and $d > 2$, the LHV polytope is also no longer confined to n -partite linear functions on \mathbf{x} .

The effect of processing on LHV correlators in the general framework of Bell tests is pointedly different from the two-measurements, two-outcome case. We hope to illustrate that there is a richness in structure in how information processing affects classical correlations and an appreciation of the connections between computation and correlators to give a handle on the effect of processing. In [7], the effect of processing can be seen to have a role in creating so-called loopholes in Bell tests and it is worthwhile exploring loopholes in the framework of data processing described here.

We also hope these results really emphasize why the $(n, 2, 2)$ setting is so special and such a powerful platform for distinguishing between classical and non-classical correlators. Even with linear Boolean pre-processing, the classical computation possible with LHV correlators is well characterized. This is still a computationally interesting setting as two inputs and two outputs at each site already enables quantum computing and the richness of structure that comes with that. Despite the fact that MBQC can be generalized to the setting with more inputs and outputs at each site, the same LHV structure does not retain the same character in this setting.

V. SUMMARY AND OPEN QUESTIONS

In this paper, we have explored the relatively under-studied multi-setting, multi-outcome Bell inequality test in the context of the computational perspective of Bell tests, which has allowed us to derive new families of non-trivial Bell inequalities. We have presented our work in terms of correlators and not the full measurement statistics of an experiment for clarity and convenience. However, we have shown that a significant amount of the structure of the space of non-signaling probability distributions are captured by this setting.

We have shown that LHV theories can be generally considered in terms of limited computational expressiveness, i.e. LHV theories can only evaluate n -partite linear functions as defined. A violation of a Bell inequality defined by these

functions implies a computational advantage relative to any LHV theory. Given this interpretation, we can construct an infinite class of non-trivial Bell inequalities. Furthermore we have explicitly constructed facet Bell inequalities that define the region of LHV correlators using the construction of LHV correlators in this paper.

Central to this paper is the incorporation of a computational perspective using modulo arithmetic into the more established picture of the convex polytopes of LHV theories. This perspective provides a broad arena to explore and, while we have given some insight, many open questions still remain. Firstly, what is the general correspondence between non- n -partite linear functions and facet Bell inequalities? There seems to be a one-to-one or a one-to-two correspondence in certain settings but this is not always the case. There are many happy coincidences that need further clarification.

A huge open question is what exactly defines the quantum region in our space of correlators? For the $(n, 2, 2)$ correlator Bell test, we know that measurements on GHZ states form the extreme points of the quantum region [10, 11]. However, little is known for other settings. At the moment finding the eigenvectors for maximal violations of a Bell inequality has not resulted in a general expression for the states that maximally violate the CGLMP inequality, if such a closed form expression even exists. In addition, more study is needed to determine the effect of pre-processing on quantum correlators.

In [7], signaling correlators were simulated with post-selection on measurement data. We have shown that the generalization of this post-selection leads to a boost in the power of LHV correlators; something that is akin to a loophole in a Bell test. However, in the more general setting, does simulating signaling correlators also boost LHV correlators? Are there more nuanced ways of processing data that does not expand the region of LHV correlators as occurs in Proposition 7?

If a violation of a Bell inequality is a resource for information processing [2–4, 6, 8], then all aspects of Bell tests need to be explored to determine how we can best exploit this resource. We have tried to hint at how the richness of structure of abstract local hidden variable theories can be captured by a simple computational picture. We hope that this work gives more insight into the cooperative relationship between quantum information and the foundations of quantum mechanics.

Acknowledgements - We would like to thank Hussain Anwar, Nicolas Brunner, Miguel Navascués and Stephen Bartlett for interesting discussions. MJH is financially supported by the EPSRC (UK) and DEB acknowledges the support of a Leverhulme Fellowship.

-
- [1] J. S. Bell, *Physics* 1, 195 (1964).
 - [2] A. Acín et al, *Phys. Rev. Lett.* 98 230501 (2007).
 - [3] S. Pironio et al, *Nature* 464, 1021 (2010).
 - [4] H. Buhrman, R. Cleve, S. Massar, R. de Wolf, *Rev. Mod. Phys.* 82, 665 (2010).; S. Pironio, *Phys. Rev. A* 68, 062102 (2003).; Č. Brukner, M. Zukowski, J.-W. Pan, and Anton Zeilinger, *Phys.*

- Rev. Lett.* 92, 127901 (2004).
- [5] R. Cleve, P. Hoyer, B. Toner, and J. Watrous, *Proc. of IEEE Conference on Computational Complexity*, pages 236–249 (2004).
- [6] J. Anders, D. E. Browne, *Phys. Rev. Lett.* 102, 050502 (2009).
- [7] M. J. Hoban and D. E. Browne, arXiv:1102.1438

- [8] M. J. Hoban et al, *New J. Physics*, 13 023014 (2011).
- [9] J. F. Clauser, M. A. Horne, A. Shimony and R. Holt, *Phys. Rev. Lett.* 23, 880 (1969).
- [10] R. F. Werner and M. M. Wolf, *Phys. Rev. A* 64, 032112 (2001).
- [11] M. Żukowski, Č. Brukner, *Phys. Rev. Lett.* 88 210401 (2002).
- [12] I. Pitowsky, *Quantum Probability - Quantum Logic, Lecture Notes in Physics* 321, Springer (1989).
- [13] A. Acin et al, *Phys. Rev. Lett.* 92, 250404 (2004); W. Son, J. Lee, and M. S. Kim, *Phys. Rev. Lett.* 96, 060406 (2006).
- [14] Daniel Collins, Nicolas Gisin, Noah Linden, Serge Massar, and Sandu Popescu, *Phys. Rev. Lett.* 88, 040404 (2002).
- [15] D. Kaszlikowski, L. C. Kwek, J.-L. Chen, M. Żukowski, and C. H. Oh, *Phys. Rev. A* 65, 032118 (2002).
- [16] K. F. Pál, and T. Vértesi, *Phys. Rev. A* 77, 042105 (2008); K. F. Pál, and T. Vértesi, *Phys. Rev. A* 79, 022120 (2009); T. Vértesi, and K. F. Pál, *Phys. Rev. A* 84, 042122 (2011).
- [17] W. Laskowski, T. Paterek, M. Żukowski, and Č. Brukner, *Phys. Rev. Lett.* 93, 200401 (2004); Marcin Wieśniak, P. Badziag, and M. Żukowski, *Phys. Rev. A* 76, 012110 (2007); Y.-C. Wu, P. Badziag, M. Wieśniak, and M. Żukowski, *Phys. Rev. A* 77, 032105 (2008).
- [18] D. Collins and N. Gisin, *J. Phys. A* 37, 1775 (2004).
- [19] T. Vértesi, S. Pironio, and N. Brunner, *Phys. Rev. Lett.* 104, 060401 (2010).
- [20] J. F. Clauser and M. A. Horne, *Phys. Rev. D* 10, 526-535 (1974).
- [21] S. Popescu, and D. Rohrlich, *Foundations of Physics* 24, 379-385 (1994).
- [22] J. Barrett et al, *Phys. Rev. A* 71, 022101 (2005).
- [23] M. L. Almeida et al, *Phys. Rev. Lett.* 104, 230404 (2010).
- [24] M. Navascués, H. Wunderlich, *Proc. Roy. Soc. Lond. A* 466, 881-890 (2009).
- [25] J. Oppenheim, S. Wehner, *Science* 330, 6007, 1072-1074 (2010).
- [26] G. Brassard et al, *Phys. Rev. Lett.* 96, 250401, (2006).
- [27] M. Pawłowski et al, *Nature* 461, 1101 (2009).
- [28] M. Froissart, *Nuovo Cimento B* 64, 241 (1981).
- [29] A. Peres, *Foundations of Physics* 29, 589-614 (1999).
- [30] J. S. Bell, *Dialectica* 39, 103 (1985).
- [31] J. Kofler, T. Paterek, and Časlav Brukner, *Phys. Rev. A* 73, 022104 (2006).
- [32] J. Barrett and N. Gisin, *Phys. Rev. Lett.* 106, 100406 (2011).
- [33] M. J. W. Hall, *Phys. Rev. A* 84, 022102 (2011).
- [34] B. Grunbaum, *Convex Polytopes (2nd edition)*, Springer (October 1, 2003).
- [35] S.-W. Lee, Y. W. Cheong, and J. Lee, *Phys. Rev. A* 76, 032108 (2007).
- [36] A. Fine, *Phys. Rev. Lett.* 48, 291 (1982).
- [37] S. Pironio, J.-D. Bancal, and V. Scarani, *J. Phys. A: Math. Theor.* 44, 065303 (2011).
- [38] G. Svetlichny, *Phys. Rev. D* 35, 3066 (1987).
- [39] E. Gawrilow, and M. Joswig, *Polytopes, combinatorics and computation (Oberwolfach, 1997)*, pp. 43-73, DMV Sem., 29, Birkhaeuser, Basel (2000).
- [40] L. Masanes, *Quant. Inf. Comput.* 3, 345 (2003).
- [41] N. D. Mermin, *Phys. Rev. Lett.* 65, 1838 (1990).
- [42] See supplemental material.
- [43] V. Paulsen, *Completely Bounded Maps and Operator Algebras*, Cambridge University Press, (2003).
- [44] D. M. Greenberger, M. A. Horne, A. Zeilinger, in *Bell's Theorem, Quantum Theory, and Conceptions edited by M. Kafatos (Kluwer Academic, Dordrecht, 1989)*.
- [45] M. Navascués, S. Pironio, and A. Acin, *New J. Phys.* 10, 073013 (2008).
- [46] A. Acin, T. Durt, N. Gisin, and J. I. Latorre, *Phys. Rev. A* 65, 052325 (2002).
- [47] J. Sturm, *SeDuMi, a MATLAB toolbox for optimization over symmetric cones*, <http://sedumi.mcmaster.ca>; J. Lofberg, *Yalmip : A toolbox for modeling and optimization in MATLAB*, <http://users.isy.liu.se/johanl/yalmip/>.
- [48] R. Raussendorf, and H. J. Briegel, *Phys. Rev. Lett.* 86, 5188-5191 (2001); H. J. Briegel, D. E. Browne, W. Dur, R. Raussendorf, M. Van den Nest, *Journal of Nature Physics* 5 1, 19-26 (2009).
- [49] D. L. Zhou, B. Zeng, Z. Xu, and C. P. Sun, *Phys. Rev. A* 68, 062303 (2003).

Appendix A: Facet Bell inequalities

In this section we list the facet Bell inequalities which belong to symmetry classes that are not explicitly described in the paper. We express the correlators $p(k|\mathbf{s})$ as the elements of a $(d-1)c^n$ vector in real space $\mathbb{R}^{(d-1)c^n}$. To simplify we describe the inequalities in terms of a vector \vec{b} that has inner product with correlator column vectors \vec{p} to form the Bell inequality, i.e. $\vec{b} \cdot \vec{p} \leq \gamma_{\mathcal{L}}$. The coefficients correspond to elements going from left-to-right as $p(1|0, 0, \dots, 0)$, $p(2|0, 0, \dots, 0)$ up to $p(d-1|c-1, c-1, \dots, c-1)$. We will use this notation in the next two sections of the appendix. For the (2, 2, 5) setting, there are 4 symmetry classes of facet Bell inequalities (excluding the class of positivity and normalization inequalities). One of these is the CGLMP inequality and the other three are given by the vectors of coefficients

$$\begin{aligned}\vec{b}_1 &= \frac{1}{2} (6, 2, 3, 4, 4, -2, 2, 1, 4, -2, 2, 1, -4, 2, -2, -1), \\ \vec{b}_2 &= (3, 1, -1, -3, 2, -1, -4, -2, 2, -1, -4, -2, -2, 1, 4, 2), \\ \vec{b}_3 &= (2, -1, 1, -2, 3, 1, -1, 2, 3, 1, -1, 2, -3, -1, 1, -2).\end{aligned}\tag{A1}$$

Denoting by \mathcal{I}_j the inequality corresponding to \vec{b}_j , \mathcal{I}_1 and the CGLMP inequality are maximally violated by the vertex corresponding to $f(\mathbf{s}) = s_1 s_2 + 1$ and \mathcal{I}_2 and \mathcal{I}_3 are maximally violated by the vertex corresponding to $f(\mathbf{s}) = 2s_1 s_2 + 1$.

The LHV bounds for all of the inequalities are listed in Tab. III with their quantum violation.

As we have seen there is a corresponding function for each of these inequalities that leads to a maximal violation. Each of these functions is a vertex in the general space of Bell correlators mentioned in the paper.

Appendix B: Generalization to non-prime number of settings

In this paper, we focused on the case where c and d are prime. However, this assumption is not needed to prove all the results. We now give a quick overview of how we describe LHV correlators in terms of polynomials on inputs for non-prime c . For non-prime number of settings, the number of inputs at each site c can be just expressed as the Cartesian product of the prime factors of c , i.e. $\mathbb{Z}_c \rightarrow \mathbb{Z}_{c_1} \times \mathbb{Z}_{c_2} \times \dots \times \mathbb{Z}_{c_q}$ where $\{c_1, c_2, \dots, c_q\}$ is the set of prime factors of c and q is

	\vec{b}
\mathcal{B}_1	2 2 1 1 2 -1 -1 -2 1 -1 -2 2 1 -2 2 1
\mathcal{B}_2	2 2 1 1 2 -1 -1 -2 1 -2 2 1 1 -1 -2 2
\mathcal{B}_3	2 2 1 1 2 -1 -2 -1 1 -2 1 2 1 -1 2 -2
\mathcal{B}_4	2 2 1 1 1 -1 2 -2 1 -2 1 2 2 -1 -2 -1
\mathcal{B}_5	2 2 1 1 1 -2 2 1 1 -1 -2 2 2 -1 -1 -2
\mathcal{B}_6	2 1 1 0 1 -1 -1 1 1 -1 -1 0 1 -1 0
\mathcal{B}_7	2 1 1 0 1 -1 -1 1 0 1 -1 0 1 -1 -1 -1
\mathcal{B}_8	2 1 1 0 0 1 -1 0 1 -1 -1 1 1 -1 -1 -1
\mathcal{B}_9	2 1 0 1 1 -1 1 -1 0 1 0 -1 1 -1 -1 -1
\mathcal{B}_{10}	2 1 0 1 0 1 0 -1 1 -1 1 -1 1 -1 -1 -1
\mathcal{B}_{11}	2 0 1 1 0 0 1 -1 1 1 -1 -1 1 -1 -1 -1
$\mathcal{C}_{c=4}^1$	1 1 0 0 1 -1 0 0 0 0 0 0 0 0 0 0
$\mathcal{C}_{c=4}^2$	1 1 0 0 0 0 0 0 1 -1 0 0 0 0 0 0
$\mathcal{C}_{c=4}^3$	1 0 1 0 0 0 0 0 1 0 -1 0 0 0 0 0

TABLE IV: The facet Bell inequality expressions that each belong to a particular symmetry class for $(2, 4, 2)$. Each row corresponds to a particular inequality belonging to a different symmetry class. Each column of \vec{b} is an element of this vector that forms an inner product with \vec{p} . The LHV upper bound for inequalities $\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3, \mathcal{B}_4$, and \mathcal{B}_5 is 8 and 4 for $\mathcal{B}_6, \mathcal{B}_7, \mathcal{B}_8, \mathcal{B}_9, \mathcal{B}_{10}$, and \mathcal{B}_{11} .

the number of prime factors in c . As a result each input s_j can be represented as a string of elements $s_j = \{t_j^1, t_j^2, \dots, t_j^q\}$ where $t_j^k \in \mathbb{Z}_{c_k}$.

For LHV theories we consider the single site maps m_j and then take their sum modulo d . For non-prime dimension then, each deterministic single site map can be written as

$$m_j = \alpha_j(\lambda) + \sum_{\substack{v \in \mathbb{Z}_c, \\ v \neq \mathbf{0}}} \beta_j^v(\lambda) \delta_v^s, \quad (\text{B1})$$

where $\alpha_j(\lambda) \in \mathbb{Z}_d$ and $\beta_j^v(\lambda) \in \mathbb{Z}_d$ dependent on the local hidden variable λ and $\mathbf{0} = \{0, 0, \dots, 0\}$. We can use the redefinition of the delta function over $\mathbb{Z}_c \rightarrow \mathbb{Z}_{c_1} \times \mathbb{Z}_{c_2} \times \dots \times \mathbb{Z}_{c_q}$ used for each prime dimension \mathbb{Z}_{c_k} . This results in the map:

$$m_j = \alpha_j(\lambda) + \sum_{\substack{v \in \mathbb{Z}_c, \\ v \neq \mathbf{0}}} \beta_j^v(\lambda) \prod_{k=1}^q \left[1 - (t_j^k - v_j^k)^{(c_k-1)} \right]_{c_k}. \quad (\text{B2})$$

One then takes the sum modulo d of all of these single site maps to obtain (a superficially complicated) map which is n -partite linear as it only consists of single-site maps. There is also no multiplication between values of t_j^k with other values of $t_{j'}^{k'}$ for $j \neq j'$. Again there are $d^{m(c-1)+1}$ possible deterministic maps for LHV theories, and their convex hull forms the LHV polytope \mathcal{L} .

The smallest value of non-prime d is 4 and so we studied the structure of the LHV polytope for the $(n, 4, 2)$ setting which has 2^7 n -partite linear functions corresponding to the vertices of the LHV polytope. There are 27,968 facet Bell inequalities for this setting found with Polymake [39] which reduces to 15 symmetry classes (14 excluding the class of normalization and

positivity inequalities). Three of these classes are forms of the CHSH inequality embedded in the larger number of inputs. For completeness, we have listed all 14 Bell inequalities in Tab. IV. We now explicitly write out one of these inequalities:

$$\mathcal{C}_{c=4}^1 = \sum_{\mathbf{s}} (-1)^{s_1 s_2} \prod_{j=1}^2 (\delta_0^{s_j} + \delta_1^{s_j}) p(\sum_{j=1}^2 m_j = 1 | \mathbf{s}) \leq 2. \quad (\text{B3})$$

which is almost exactly the same as $\mathcal{C}_{c=3}^1$. The other two inequalities, $\mathcal{C}_{c=3}^2$ and $\mathcal{C}_{c=3}^3$ are similar to this inequality except with altered delta functions for $\mathcal{C}_{c=3}^2$:

$$\prod_{j=1}^2 (\delta_0^{s_j} + \delta_1^{s_j}) \rightarrow (\delta_0^{s_1} + \delta_2^{s_1})(\delta_0^{s_2} + \delta_1^{s_2}), \quad (\text{B4})$$

and for $\mathcal{C}_{c=3}^3$:

$$\prod_{j=1}^2 (\delta_0^{s_j} + \delta_1^{s_j}) \rightarrow (\delta_0^{s_1} + \delta_2^{s_1})(\delta_0^{s_2} + \delta_2^{s_2}). \quad (\text{B5})$$

Throughout the whole of the paper, we have assumed that d has been prime. Indeed, this fact has been vital in the proof of Theorem 2, so this proof fails for non-prime d . However, for other aspects of the methods in this paper, a non-prime d has little consequence for our construction of correlators. There are two approaches to considering non-prime d : the most obvious approach is just to think of representing outcomes as a string of digits from prime number registers as with the description of settings above; secondly, instead of having the elements $f(\mathbf{s})$ of a function as elements of a vector in a vector space over \mathbb{Z}_d for prime d , we now have elements of a *module* over \mathbb{Z}_d ; a generalization of a vector space. Since we only ever take addition of functions represented in a Kronecker delta function basis to generate all functions, all of our methods generalize to these modules.

From the aspect of functions, the only difference for non-prime d is the fact that $f(y)$ in (2) can now take values from a non-prime register \mathbb{Z}_d . The derivation of all functions and n -partite linear functions generalizes naturally. For $(2, 2, 4)$ with the smallest non-prime $d = 4$, we present the facet Bell inequalities. There are four symmetry classes for this setting of which one is the class of positivity and normalization inequalities. One of these three non-trivial symmetry classes is formed by the CGLMP inequality. The second is another generalization of the CHSH inequality and is of the form:

$$\sum_{\mathbf{s}} (-1)^{s_1 s_2} \left[p(\sum_{j=1}^2 m_j = 1 | \mathbf{s}) + p(\sum_{j=1}^2 m_j = 3 | \mathbf{s}) \right] \leq 2, \quad (\text{B6})$$

which is essentially the CHSH inequality if each party groups their outcomes m_j into modulo 2 terms. In other words, since $1 \bmod 2$ is equal to $3 \bmod 2$, the above inequality is equivalent to the CHSH inequality if each party just maps from mod 4 arithmetic to mod 2.

The third and final symmetry class is generated by the following inequality (expressed in the notation described earlier):

$$(1, 2, 1, 1, 2, 1, 1, 2, 1, -1, -2, -1) \cdot \vec{p} \leq 4. \quad (\text{B7})$$

The deterministic map or function $f(\mathbf{s})$ that achieves the max-

imal upper bound 6 of this inequality for all correlators in \mathcal{P} is $f(\mathbf{s}) = 2s_1s_2 + 2$. It is worth noting that this can be constructed by adding $\sum_{\mathbf{s}} 2(-1)^{s_1s_2}p(2|\mathbf{s})$ to the previous inequality.